AU/ACSC/088/1998-04

AIR COMMAND AND STAFF COLLEGE

AIR UNIVERSITY

AIR FORCE DEVELOPMENT TEST AND EVALUATION UNDER ACQUISITION REFORM: SIX FLIGHT TEST PROCESS CONCEPTS

by

JoAnn L. Erno, Major, USAF

A Research Report Submitted to the Faculty

In Partial Fulfillment of the Graduation Requirements

Advisor: Major Fred L. Warren

Maxwell Air Force Base, Alabama

April 1998

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (FROM - TO)
01-04-1998	Thesis	xx-xx-1998 to xx-xx-1998

4. TITLE AND SUBTITLE
Air Force Development Test and Evaluation Under Acquisition Reform: Six Flight Test Process Concepts
Unclassified

5a. CONTRACT NUMBER
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Emo, JoAnn L. ;

5d. PROJECT NUMBER
5e. TASK NUMBER
5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME AND ADDRESS
Air Command and Staff College
Maxwell AFB, AL36112

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME AND ADDRESS
,

10. SPONSOR/MONITOR'S ACRONYM(S)
11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
APUBLIC RELEASE
,

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Acquisition Reform changed government involvement in system acquisition with transition to best commercial practices. Shortened acquisition timelines, reduced budgets and increased contractor test responsibilities intensified several issues. These issues are test planning and execution capabilities, joint contractor and government test efficiency, and both government and industry test experience levels. Today?s DT&E methods are insufficient to successfully overcome increased system complexity, decreased test budgets, and reduced manpower challenges of the future. To meet these challenges, this project presents six test recommendations which enhance DT&E contributions to the acquisition process. The six recommendations are: 1. Develop a single commercial industry standard governing military aviation testing. 2. Emphasize greater simulation configuration control and technological simulation advances, with software test and avionics integration simulation the highest priorities. 3. Refine the Integrated Test Team Concept to combine contractor, development, and operational test requirements in addition to innovative test program teaming. 4. Streamline development and operational test data requirements and allow contractor and development test personnel to supplement operational test teams. 5. Automate test planning and execution functions to reduce manpower investments. 6. Modify the Federal Aviation Administration Designated Engineering Representative concept to increase contractor test autonomy.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19. NAME OF RESPONSIBLE PERSON
			Public Release	57	Fenster, Lynn lfenster@dtic.mil
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER International Area Code Area Code Telephone Number 703767-9007 DSN 427-9007

Disclaimer

Contents

Page

Tables

Page

Preface

My interest in Development Test and Evaluation (DT&E) under Acquisition Reform stems from personal experience working as both a DT&E and Operational Test and Evaluation (OT&E) Engineer, and most recently a C-130 Program Manager. After spending ten years involved with development and operational test programs, Acquisition Reform seemed to lessen the value of a government role I felt was very important. Since Acquisition Reform significantly changed all the operating norms which the test community had spent decades building, I assumed the changes were for the worst. Test personnel are trained to be conservative and do not accept change readily. I set out on this project hopeful to learn that the test and evaluation community had adjusted to a new paradigm and was adapting to the commercial practices environment.

My analysis results were surprising with several successes reported in test procedural and teaming arrangements which I thought would take several more years before effects would be seen. I also expected my analysis to demonstrate government development test and evaluation was absolutely necessary. In this case, my analysis results were inconclusive because complete program failures, due to the lack of government conducted DT&E, were not reported in the analysis database. Conclusions which were supported include: DT&E was necessary, organization bias must be removed, and well trained test personnel were required to accomplish test programs efficiently. From personal experience, that sounds easier than it actually is, especially considering bias is

difficult to ascertain and well trained personnel are very expensive. I anticipated concrete results showing government test personnel membership yielded significant benefits later in program execution by inclusion of this important expertise. In this case, the conclusion supported is the government has a pool of well trained test personnel who are regularly accomplishing the test mission under difficult conditions. That doesn't mean that the government is the only organization that can efficiently execute a test program, it means the government has test trained personnel who understand the military environment *at this moment*. Overall, knowledge of military systems and how we use our equipment are crucial factors in future reform, and without it, products will always fall short of their potential. Our greatest development test contribution in the future will be information transfer accomplished more efficiently and much faster.

Finally, I would like to express my sincere appreciation to Major Fred Warren for his support and assistance in this research project. His patience and input helped significantly in producing a better research paper. Also, thank you to Mr. Jan Howell, whose PACER CRAG information provided insight to a commercial acquisition pilot program. Finally, to Major Neil Erno and my son Michael two special thank yous for your patience and wonderful smiles.

Abstract

Acquisition Reform changed government involvement in system acquisition with transition to best commercial practices. Shortened acquisition timelines, reduced budgets and increased contractor test responsibilities intensified several issues. These issues are test planning and execution capabilities, joint contractor and government test efficiency, and both government and industry test experience levels. Today's DT&E methods are insufficient to successfully overcome increased system complexity, decreased test budgets, and reduced manpower challenges of the future. To meet these challenges, this project presents six test recommendations which enhance DT&E contributions to the acquisition process.

The six recommendations are:

1. Develop a single commercial industry standard governing military aviation testing.
2. Emphasize greater simulation configuration control and technological simulation advances, with software test and avionics integration simulation the highest priorities.
3. Refine the Integrated Test Team Concept to combine contractor, development, and operational test requirements in addition to innovative test program teaming.
4. Streamline development and operational test data requirements and allow contractor and development test personnel to supplement operational test teams.
5. Automate test planning and execution functions to reduce manpower investments.
6. Modify the Federal Aviation Administration Designated Engineering Representative concept to increase contractor test autonomy.

Chapter 1

Introduction and Analysis

> *The military advantage goes to the nation who has the best cycle time to capture technologies that are commercially available; incorporate them in weapon systems; and get them fielded first.*
>
> —Honorable Paul G. Kaminski
> Under Secretary of Defense for Acquisition and Technology, 1995[1]

Air Force Development Test and Evaluation (DT&E) is at a crossroads. Acquisition Reform reduced government involvement in program acquisition with transition to best commercial practices. The shift to commercial practices reverberated change in every aspect of government acquisition including Air Force DT&E. Shortened acquisition timelines, reduced budgets and increased contractor test responsibilities intensified several issues. These issues are test planning and execution capabilities, joint contractor and government test efficiency, and both government and industry test experience levels. Today's DT&E methods are insufficient to successfully overcome increased system complexity, decreased test budgets, and reduced manpower challenges of the future. To meet these challenges, this project presents six recommendations which enhance DT&E contributions to the acquisition process.

The lessons of recent acquisition programs provide a framework to launch the next level of DT&E reform. The proposed recommendations address improvement efforts and, in several instances, are nontraditional DT&E approaches. However, the intent is to

show there are many alternatives to further refining DT&E performance. Irrespective of which organization performs testing, the goals of test programs are the same: provide the best product by quickly identifying the best and worst aspects of a system, in the most cost efficient manner, and while there is still time to improve the product.

Acquisition Reform Background

To conduct a lesson learned analysis, reform measures specifically affecting Air Force DT&E need identification. Acquisition reform's primary goals include shortened acquisition timelines and less Air Force oversight while increasing management insight.[2] Although Acquisition Reform encompasses many initiatives, three reforms are particularly applicable to analysis. The first two initiatives originate within the Air Force and apply only to Air Force processes. The third initiative affects all services acquisition programs.

Reduction of time from requirement definition to contract award is a reform initiative which streamlines the contract decision making process.[3] Awarding contracts quickly shortens the acquisition cycle and gets products to users faster. However, shorter decision timelines, also means less time to plan and analyze available test data before contract award. A great challenge remains for the government. To reduce acquisition timelines, test and evaluation planning needs to provide sufficient information to assess program risk in less time and with fewer resources.

System Program Office (SPO) manning reduction is a second reform initiative with DT&E impact.[4] Following classified special access programs manning requirements, decreased SPO manning also reduces the size of SPO test teams. Increased

team membership diversity, with broad backgrounds and experience levels, is an effort to facilitate better government choices with fewer personnel.

Finally, the government largely eliminated military specifications and standards, in 1993, as government design and performance requirements.[5] Increased program acquisition speed was the intended result of less intrusive government involvement in program development and production. In the past, DT&E provided the SPO with necessary confirmation that a system met the government standards in addition to contractual requirements. The extensive reduction in specifications and standards removed a primary DT&E program role. Best commercial practices are now the governing guidance. However, program offices must still manage program acquisition risk. To meet the changed test environment, DT&E roles evolved to meet new test challenges.

Development Test and Evaluation Background

After 1993, DT&E began its transformation. Although the classic development specification compliance test role disappeared with the extensive reduction in specification and standards, today DT&E remains integral to the acquisition process. Former Secretary of Defense William Perry defines test and evaluation's role in risk reduction:

> During the early phases of development, T&E is conducted to demonstrate the feasibility of conceptual approaches, evaluate design risk, identify design alternatives, compare and analyze trade-offs, and estimate satisfaction of operational requirements.…The T&E process provides information to: developers for identifying and resolving technical difficulties; decision-makers responsible for procuring a new system and for the best use of limited resources; and to operational users for refining requirements and supporting development of effective tactics, doctrine, and procedures.[6]

One additional DT&E requirement remains that is not apparent in the previous T&E definition. The developing agency is required to formally certify a system is ready for dedicated Operational Test and Evaluation (OT&E).[7] OT&E means:

> [T]he field test, under realistic combat conditions, of any item of (or key component of) weapons, equipment, or munitions for the purpose of determining the effectiveness and suitability of the weapons, equipment, or munitions for use in combat by typical military users; and the evaluation of the results of such test.[8]

Completing operational certification processes requires adequate information on a system's ability to meet user's requirements. Since acquisition reform shortened timelines and reduced manpower, government DT&E organizations utilize commercial test results whenever possible to reduce independent test requirements.[9] Air Force DT&E, therefore, has a two fold purpose. The first purpose is to provide enough information to reduce program risk, in less time and with less resources, and the second is to facilitate OT&E certification.

Testing required to satisfy both risk reduction and operational certification generates conflicting program goals. New acquisition policies require minimum government involvement in a program, yet operational test certification entails proof of a systems suitability before operation testing begins. There is a balance between minimum involvement and adequate system performance information that programs must strike. Lessons from previous programs prove insightful in attaining this balance. Analysis of these lessons reveals problem areas for additional improvement and indicates a priority of additional changes needed to continue reform. However, analysis requires an assessment method to ensure conclusions are appropriate.

Assessment Methods

Three information sources provided a well rounded assessment. A primary test program lesson database provided quantitative data. Government best practices from three test programs, representative of both the Air Force and Navy, added support information. Finally, a complete test program report provided comparison and contrasting information.

The *Defense Acquisition Deskbook* (DAD) is a quarterly electronic resource used by acquisition personnel within the Department of Defense (DOD). DAD records program lessons from all functional specialties and services in a searchable database. The lessons cover a wide cross section of complete and in-progress acquisition programs. Drawing upon those lessons, classifying the primary causes is the next analysis step. **Test capabilities, test processes, and test personnel are the three primary categories used to sort program lessons**.

To further define analysis, each primary topic includes two **subcategories. Test planning** and **test execution** subcategories divide the test capabilities area**. Test redundancy** and **limitations to streamlining** refine the process topic. Finally, personnel topic subcategories are **test training** and **team membership**. Appendix A contains specific definitions of each subcategory used to classify lessons learned.

Lessons classification was primarily a subjective evaluation. The recommendation content, lesson learned description, and background information, in order of importance, defined the subcategory classifications. If the lesson learned category remained unclear, a subjective evaluation of the most likely lesson cause, combined with a comparison to similar lessons descriptions, provided the final lesson classification.

Best program practices from two Air Force and one Navy flight test program added analysis depth. Program specific recommendations from the Air Force Joint Primary Aircraft Training System supplied a comparison source of program successes.[10] The Air Force Joint Direct Attack Munition was the second program used to validate analysis results.[11] The Navy F/A-18E/F supplied an additional service program contrasting information.

Finally, the KC-135 PACER CRAG test program added additional balance to lesson analysis. As a commercial acquisition program, PACER CRAG provided specific lessons learned representative of system upgrade modernization programs.[12] PACER CRAG's test program relied on commercial industry best practices. Lessons from this program were useful to evaluate commercial test capabilities.

Analysis Results

Data analysis used 60 lessons reported in the DAD. Table 1 presents the overall results from the classifications. For additional detail, Appendix B lists the 60 lessons and individual classifications. During analysis, two data limitations affected subcategory determinations. Test lessons did not include team composition and experience level information. Classifying lessons into the test training and team membership subcategories was difficult. Lesson assignment in these categories occurred only when the lesson cause was likely due to training deficiencies or inefficient team membership, and another cause was not identified. Therefore, the number of lessons assigned to team training and membership subcategories may too low. Secondly, source selection test planning capabilities also were unknown. In this case, an assumption was necessary. Unless another category was obvious, test planning subcategory assignment was

automatic for lessons learned during source selection. Thus, the number of lessons assigned to the test planning subcategory may be too high.

Table 1. Analysis Results - Lessons Classifications

Lesson Learned Subcategories	Number of Lessons	Percentage
Capabilities - Test Planning	17	28.3%
Capabilities - Test Execution	20	33.3%
Processes - Test Redundancy	15	25.0%
Processes - Limitations to Streamlining	2	3.3%
Personnel - Training	4	6.7%
Personnel - Team Membership	2	3.3%
Totals (Percentages rounded to nearest tenth.)	60	99.9%

Source: *Defense Acquisition Deskbook.* Version 2.2. (Wright-Patterson Air Force Base, Ohio, ODUSD(AR and OUSD(A&T)) API Deskbook Joint Program Office. 15 December 1997).

Improvement Challenges

Examining data results, three questions arise. Data analysis showed test planning and execution were the greatest sources of lessons. Therefore, the first question is, "Can we improve flight test planning and execution?" Test redundancy and test streamlining were the second greatest sources of problems, raising the question, "How will government continue to minimize test redundancy and what process improvements will further streamline test programs?" Finally, adequate and experienced manpower, and more efficient use of that manpower, are the third improvement areas. Accordingly, "What alternative approaches will maximize use of trained and experienced test personnel?" In the next three chapters, analysis results and additional views present options addressing each question. The recommendations proposed are only a subset of many possible answers. However, they do show that advances in DT&E methods are achievable and necessary.

Notes

[1] Paul G. Kaminski, “Reinventing DOD Test & Evaluation,” 3 October 1995, n.p. [no pagination]; on-line, Internet, 17 September 1997, available from http://www.acq.osd.mil/te/speeches/kaminski/itea.doc.

[2] William Perry, “Mandate for Change,” briefing presented to the House Armed Services Committee and the Governmental Affairs Committee, February 1994, 11, on-line, Internet, 9 March 1998, available from http:// www.acq.osd.mil/ar/#library/ mand24.pdf.

[3] *Lightning Bolt Update #14*, SAF/AQ, Pentagon, VA, 6 August 1997, n.p.; on-line, Internet, 21 January 1998, available from http://www.safaq.hq.af.mil acq_ref/bolts/ update14.html. Lightning Bolt 10.

[4] *Lightning Bolt Update #14*, SAF/AQ, Pentagon, VA, 6 August 1997, n.p.; on-line, Internet, 21 January 1998, available from http://www.safaq.hq.af.mil/acq_ref/bolts/ update14.html. Lightning Bolt 3.

[5] *Federal Participation in the Development and Use of Voluntary Standards,* OMB Circular A-119, 20 October 1993, 9, on-line, Internet, 4 February 1998, available from http://www.afam.wpafb.af.mil.

[6] “Key Test Issues,” *Defense Acquisition Deskbook,* Version 2.2, Wright-Patterson Air Force Base, Ohio, ODUSD(AR and OUSD(A&T) API Deskbook Joint Program Office, 15 December 1997, n.p.; on-line, Internet, 4 February 1998, available from http://www.afam.wpafb.af.mil.

[7] DOD 5000.2-R *Mandatory Procedures for MDAPs and MAIS Acquisition Programs,* Change 2, 6 October 1997, n.p.; on-line, Internet, 4 February 1998, available from http://www.afam.wpafb.af.mil. Paragraph 3.4.3.

[8] *Title 10 — 10 Uniform Service Code; Armed Forces; Subtitle A — General Military Law Sec. 139 — Director of Operational Test and Evaluation;* n.d., n.p.; on-line, Internet, 22 October 1997, available from http://www.afam.af.mil, Section 139 subparagraphs a2A ii and iii.

[9] Air Force Instruction (AFI) 99-101 *Development Test and Evaluation,* 1 November 1996, 10.

[10] “Joint Primary Aircraft Training System,” *Acquisition Reform Stories,* n.d., n.p.; on-line, Internet, 15 October 1997, available from http://www.safaq.hq.af.mil/acq-ref/stories/stories.htm.

[11] “Air Force Joint Directed Attack Munition,” n.d., n.p.; on-line, Internet, 24 September 1997, available from http://www.arnet.gov/References/Best_Pract_Docs/ AFJDAM.html.

[12] Michael E. Bonner, and Michael J.Scheall, *PACER CRAG C/KC-135 Avionics Modernization Evaluation,* AFFTC-TR-97-01, Air Force Flight Test Center., Edwards Air Force Base, California. May 1997.

Chapter 2

DT&E in a New Era: Meeting the Demand

> *…we face a dual-faceted challenge. On one hand we are working within the decreasing modernization budgets, including less resources for engineering and test as well as less investment dollars for enhancing of SE (System Engineering) and T&E capabilities. On the other hand, our Department's strategy increasingly relies on the fielding of fewer, but more capable systems—systems which are inherently more sophisticated and complex - with greater technical challenges to manage and which stress our current test and engineering capabilities.*
>
> —Dr. Patricia Sanders
> Director Test, Systems Engineering & Evaluation[1]

Flight test planning and execution are resource-intensive activities. Acquisition reform did not change the complexity of test planning and execution, but it did shorten timelines and reduced government team size. Lessons are expected in adjustments to faster contract awards and smaller teams. Indeed, the analysis results confirm the largest lessons subcategories were test planning and execution. However, the reasons for these lessons were unexpected. Test planning guidance and simulation are the future test capability challenges.

Performance Standards and Reliance on Best Commercial Practices

Test planning begins in the formation of the acquisition strategy and continues throughout the acquisition program. In the lessons analysis, test planning included test data gathering methods and test conduct philosophy. Seventeen lessons identified test

data requirements during source selection and commercial marketing practices as areas for concern. Source selection decisions, which relied on commercial test results, resulted in later test planning problems and additional testing to verify commercial tests. The reliance upon commercial test results shortened the contract award timeline but did not necessarily provide a product faster.

Commercial Test Standards and Best Practices for the Military Environment

Commercial industry's willingness to present adequate and accurate test data, and the government's determination to require this information in source selection, are the first areas for improvement. With a few exceptions, the government no longer requires adherence to military standards and specifications. Replacing these standards was reliance on commercial guidance to govern contractor methods. Reliance on commercial guidance assumes that the guidance exists either formally or informally. However, commercial guidance governing test conduct and the military environment was not found during research. Little information defining military utility or typical flight environments was available. The lack of commercial guidance formulating consistent test policies in the military environment complicated the source selection processes in two ways.

First, source selection decisions were made with little information on integrated system compatibility or well defined system pass/fail criteria. In addition, program offices relied on market research and commercially-provided test data to base contract awards.[2] However, reliance on commercially-provided test results was inadequate for many systems. Several lessons cited overly optimistic predictions of how well a system would perform in the military environment. The DOD Inspector General lessons learned illustrate the need for guidelines governing optimistic proposals.

> In the contract request-for-proposal process, the buying organizations indicated that commercial suppliers often exaggerate product performance and operating capabilities to obtain competitive advantages. FAR [Federal Acquisition Regulations] 12.205 cautions buying organizations to review product literature to determine its adequacy for purposes of evaluation through market research to avoid such problems.[3]

Following commercial practices, government teams assessed military utility using marketing information. However, due to the competitive commercial environment, there was little incentive for industry to provide information that would discount their claims. Additionally, in commercial acquisition programs, FAR 12.221 restricts data requests to data normally commercially available.[4] Thus, for products with a commercial market, source selection teams made test planning decisions with incomplete system performance information.

The military as the primary product market created a second source selection complication. In this case, initial development test planning did not address the changed government program involvement. The contractor and government responsibilities were not clear. Several programs pointed to inadequate definition of responsibilities in the initial contracting phases leading to later test problems. Contract modifications were the only method to resolve the issues which added time and cost to the program. In one case, three contract revisions occurred to correct acceptance test documents.[5]

Commercial Test Standard Challenge

The program office must transmit to offerors the government's test requirements. However, understanding government test requirements is the responsibility of commercial industry. Good source selection and test planning decisions are dependent upon compete information regarding system utility. Commercially-maintained standards and best practices are one way to reduce test planning problems. The PACER CRAG

program underscores the need to continue progress in test planning. The avionics upgrade test effort grew by 400% from initial contract to completion of the test program.[6] Commercial and military system integration difficulties, led to a commercial test effort significantly more complex than originally planned.

An industry wide standard governing military testing and reporting will streamline source selection and improve government and contractor test team relationships. Reliance on commercial best practices is dependent upon commercial industry's willingness to set those standards. The military's role in providing test standards has diminished; the challenge lies with commercial industry to provide comprehensive test guidance for military aviation.

Simulation Needs

Multiple Test Requirements

Following test planning, efficient test execution also relies on accurate and complete data. Simulation is one tool used to provide additional test data while also reducing test requirements. In lesson analysis, simulation capability exhibited several problems. First, simulation experienced technical problems in software intensive avionics systems and new system integration programs. Two programs reported software problems found during flight test which simulation did not adequately predict.[7] New software versions required testing to ensure the changes worked in addition to regression testing which verifies the fixed software did not degrade other system functions. Increased test requirements and program delays are the results. AFMC software regression testing lesson documents program schedules impact:

> This failure to take a global independent look at required regression testing for software maintenance action can result in later system surprises and further deteriorated schedules as a result of undetected problems added by the maintenance action. It is necessary to perform regression testing when software changes are made to ensure the software still meets the performance and design requirements.[8]

A simulation capability adequate to find more software problems before a test program begins and lessen regression test requirements is a technical challenge to both the government and industry.

The second simulation concern was configuration lag between the simulator and the test article. The F-22 Safety Board documents the simulation configuration problem in their findings and recommendations for the F-22 simulator. Specifically mentioned was the difficulty keeping standard software configurations in the simulators.[9] In this case, simulation was unable to support flight test because software versions between the simulator and the test aircraft were not the same. Simulation is a tool supporting test execution. However, lessons show simulation effectiveness was less than optimum.

A Simulation Technical Challenge

Simulation was meant to reduce the amount of flight test. Unfortunately, the inability to find software problems early, and later simulator configuration problems, reduced simulation's effectiveness. Future systems will continue increasing in software complexity and the introduction of more commercial systems will create continued software compatibility issues. However, timeline and cost reduction measures will certainly apply pressure to minimize flight test programs. Simulation continues to grow in significance, and management emphasis to provide the necessary capability must grow accordingly.

The DOD recognizes simulation's importance in future test and evaluation. In 1995, Dr. Kaminski, Under Secretary of Defense for Acquisition and Technology, relayed the emphasis on simulation:

> The Departments' senior leadership is strongly committed to greater use of modeling and simulation, especially models that incorporate real physical underpinnings. With such models, we can actually eliminate certain tests and focus test resources on the areas where our understanding is less. In many cases, we should be conducting tests to validate our models and simulations.[10]

Simulation is used to reduce flight test requirements. However, there are limitations to how much flight test may be trimmed. Simulation's test contributions have led to beliefs that simulation will replace flight test. The thought that simulation is more than a flight test reduction tool proved false and led to more and not fewer flight test requirements.[11] Mr. Robert Myers, Northrop Grumman Vice President of Test Operations, cautions on simulation:

> People think we can progress faster than we really can, based on modeling, simulation and laboratory testing,…The question is, how much do you reap when you're in the flight test phase? You don't get quite as much [gain] as predictions would indicate.[12]

Two simulation advantages are reduced regression test and flight test requirements. To increase simulation contributions in these areas requires higher priority on simulation improvements for software intensive systems and simulation configuration control. In addition, as newer systems increase in software complexity, improved simulation technology is a necessity.

The advantages of improved simulation are faster and cheaper acquisitions. Improved simulation also requires early management emphasis and priority for monetary resources. Both government and industry must support the cost trade-off decisions to

advance simulation capability. Future programs which place a priority on simulation will enjoy efficient test phases later in the acquisition program.

Notes

[1] Dr. Patricia Sanders, Statement by the Director Test, Systems Engineering & Evaluation to the Subcommittee on Military Research & Development of the House National Security Committee, 27 February 1997, n.p.; on-line, Internet, 15 October 1997, available from http://www.acq.osd.mil/te/speeches/anders/testimony.htm.

[2] *Defense Acquisition Deskbook,* Version 2.2, Wright-Patterson Air Force Base, Ohio, ODUSD(AR and OUSD(A&T) API Deskbook Joint Program Office, 15 December 1997, n.p.; on-line, Internet, 12 January 1998, available from http://www.afam.wpafb.af.mil, Lessons from the DOD IG-Test and Evaluation of Commercial Items.

[3] Ibid., Lessons from the DOD IG-Test and Evaluation of Commercial Items.

[4] FAR Part 12, *Acquisition of Commercial Items*, 9 February 1998, 5, Section 12.211 - Technical Data states "…the Government shall acquire only the technical data and the rights in that data customarily provided to the public with a commercial item or process."

[5] *Defense Acquisition Deskbook,* Version 2.2, Wright-Patterson Air Force Base, Ohio, ODUSD(AR and OUSD(A&T) API Deskbook Joint Program Office, 15 December 1997, n.p.; on-line, Internet, 12 January 1998, available from http://www.afam.wpafb.af.mil, AFMC - T&E Lessons Learned -RFP Prep/Source Selections, 10. Development and Approval of Acceptance Test Procedures.

[6] Jan Howell, "PACER CRAG — A COTS/NDI Lightning Bolt Program," briefing, Air Force Flight Test Center, Edwards AFB, CA., n.d.

[7] *Defense Acquisition Deskbook,* Version 2.2, Wright-Patterson Air Force Base, Ohio, ODUSD(AR and OUSD(A&T) API Deskbook Joint Program Office, 15 December 1997, n.p.; on-line, Internet, 12 January 1998, available from http://www.afam.wpafb.af.mil, AFMC - T&E Lessons Learned for Evaluating System Test Data, 4. Random Software Test Failures, and 5. Flight Test Data Analysis which stated: "Vehicle and software design problems were noticed only in overt failures. Early indications of such problems were overlooked or not addressed."

[8] Ibid., AFMC - T&E Lessons Learned -RFP Prep/Source Selections, 5. Software Regression Testing.

[9] Ibid., AFMC - Lessons Learned for Detailed Test Planning, 8. Configuration Control Deficiency at Simulation Facility.

[10] Paul G. Kaminski, "Reinventing DOD Test & Evaluation," 3 October 1995, n.p. [no pagination]; on-line, Internet, 17 September 1997, available from http://www.acq.osd.mil/te/speeches/kaminski/itea.doc.

[11] William B. Scott, "New Global Pressures Reshape Flight Testing," *Aviation Week & Space Technology,* 12 June 1995, Vol. 142 Issue 24, 62. CD-ROM. UMI-Proquest, 15 October 1997.

[12] Ibid.

Chapter 3

An Era of Jointness

It is clear that a bulk of our money is spent on manpower—at least 54% (sic) for government personnel and an additional 15 percent for contractor services.…Fundamentally, manpower is the high leverage point - if a reinvention idea does not reduce required manpower, it will not have a significant impact..

—Dr. Patricia Sanders
Director Test, Systems Engineering & Evaluation[1]

Test and evaluation requires an efficient execution concept to maximize use of limited manpower and test resources. Chapter 2 showed how a test team's ability to quickly and efficiently accomplish a test program is directly related to the quality of its test planning and its access to adequate simulation. After test capabilities, the test process category was the next largest area generating lessons learned. Multiple test organizations each had independent test teams and test requirements. In addition, limitations to test streamlining exist and restrain test integration. One possible concept, which addresses both test efficiency and team manning, is an advanced integrated test team.

Integrated Test Team Concept

Test Team Membership

Contractor, development, and operational test organizations each have independent test requirements. Since the relationship between each test organization changed with acquisition reform, all three organizations are now very dependent upon each other to accomplish test programs. Program lessons showed test team interdependency did not lessen test requirements nor create an efficient test team structure. Contractor test team membership, particularly smaller systems integration programs, lacked test-trained and test-experienced personnel. Secondly, SPO downsizing also reduced development test team size below sustainable levels. The PACER CRAG program also experienced stressed development test teams.[2] In the PACER CRAG example, optimistic test schedules drove initial development test manpower levels. After system problems were found, the test program lengthened and the development test team struggled to support the program.[3] Overall, test requirements and test team manning were mismatched. Sharing manpower among contractor, development and operational test teams would address the resource shortfalls.

Cross-organization test teaming is a concept with documented results. The combined contractor, development, and operational test concept has worked well over the last 20 years.[4] In addition, initial results from similar teaming arrangements are positive. The results were decreased manpower requirements, and reduced test. The F/A-18 E/F program reported their success using an integrated team approach:

> The Integrated Test Team (ITT) is a departure from the traditional test team assembled to support an Engineering and Manufacturing Development (EM&D) (sic) effort. From its inception in August 1995,

> this team has fully integrated government and industry developmental and operational test communities to make the best, most efficient use of time and assets in performing the most comprehensive testing possible. … Joint execution of the test plan permitted data sharing, while further reducing the need for duplicitous (sic) test, dramatically reducing cost, schedule and assets required.[5]

However, the integration of teams must go further to include a completely shared manpower pool to support all test requirements. The team solution must use the test teams' abilities in an even more efficient and deliberate manner. The advanced team concept includes using contractor personnel testing systems to fulfill operational data requirements. Lessons support contractor operational test involvement. One lesson cited problem documentation which did not flow to the operational test team:

> The failure to monitor and document potential or known problems during contractor engineering tests and development test and evaluation (DT&E) places an increased burden on the operational test team to effectively identify and document problems which continue during the relatively short time when IOT&E [Initial Operational Test and Evaluation] is conducted.[6]

Contractor membership on the operational test team would streamline information flow while also increasing test efficiency.

Streamlining test programs requires well organized test teams. Although the integrated test team arrangement is a well documented concept, there is still room for additional improvement.

Further Integration of Contractor and Government Teams

Test team integration continues with multiple programs building upon previous lessons learned. The Navy F/A-18 E/F accomplished integration of contractor and government teams by clearly defining each organization's roles and ensuring overlapping responsibilities were kept to a minimum:

> The ITT conducts all contractually required functional, system, and air vehicle flight tests. The Contractor led ITT is responsible for the conduct of the development flight test program.…The contractor remains accountable for all contractual efforts.…The government members of the ITT are primarily responsible for identifying and incorporating unique requirements into the Master Test Plans and TWDs (Test Work Descriptions).…Government members of the ITT review generated data for adequacy, evaluate and analyze data to support flight test, assess mission effectiveness, and ensure that government management in informed on program status and issues.…Government and Contractor pilots participate in all phases of the E&MD (sic) fight test program.[7]

Well-defined roles led to efficient team arrangements, but, several programs reported a problem which limited further team integration. Test team integration success depended upon the contractor's testing philosophy. Discussed in Chapter 2, many programs reported optimistic contractor test planning in initial proposals. Later, the legacies of optimistic planning were under-manned test teams and inefficient test processes. In the PACER CRAG example, initial test plans increased from six flights in three weeks to 29 flights in 11 weeks leaving the original test teams greatly under-resourced.[8] Government test teams were under-staffed to accomplish their own test responsibilities reducing their ability to support the combined test team. Optimistic contractor test plans leads to less-than-satisfactory government test team participation. Integration of government and development contractor test teams must continue but manning must be based on conservative test schedules.

There are also other limitations to test team integration. For example, OT&E has independent test team requirements mandated by law. Discussed next are the origins of these limits and their impact on further test team integration.

Policy, Laws, and Guidance

Limits to Test Streamlining

OT&E requirements originate in laws and policy that are difficult to change. Title 10 *Uniform Service Code*, Sections 139 and 2399 govern operational test independence.[9] Independent reporting requirements result in separate teams and dedicated operational testing. The Mission Needs Statement (MNS) and the Operational Requirements Document (ORD) are the two documents used to assess system military suitability. In addition, DT&E uses the same documents to complete an operational test certification.

Stated in DoD5000.2-R, DT&E organizations have a responsibility to provide an operational readiness recommendation:

> The developing agency shall prepare a DT&E Report, and formally certify that the system is ready for the next dedicated phase of operational test and evaluation…[10]

Additionally, Air Force Instruction 99-101, *Development Test and Evaluation,* expands upon the certification requirement.

> IPT (Integrated Product Team) personnel must plan to conduct government DT&E in as operationally relevant an environment as possible without compromising the goals of DT&E….Development testers will evaluate the system against user requirements (as stated in the ORD) as well as system specifications:[11]

Guidance directing operational testing originates before acquisition reform changed DT&E roles. The recent DT&E emphasis leaves duplicate requirements in both the development and operational test areas. OT&E independent assessment requires an independent test team with separate test requirements. In addition, operational testing requires a production-representative test article.[12] Test articles appropriate for operational testing have finished a development test program. Since OT&E requires

operational test certification, development testing of operational readiness is also required to complete certification. The result is two test teams with duplicated test requirements.

The development and operational test independence issue is contentious. Lessons researched supported continuation of both DT&E and OT&E. In addition, consensus was also apparent on two issues: DT&E is necessary even in a commercial acquisition situation, and OT&E is needed earlier in the acquisition cycle.[13] Integration of operational test personnel in the SPO team as early as source selection is one option, but duplicate manpower requirements still remain and actually increase with the addition of another test experienced person representing operational test interests. A second proposal involves changing operational test independence requirements, which means changing the laws. The independence of OT&E is not argued in this paper. However, an advanced integrated team concept, allowing triple use of contractor, development, and operational test personnel, is one possibility which answers all three test requirements.

Two Test Policy Changes

Before an advanced team concept can function, two issues that affect development and operational test conduct also need reform. The first issue is fulfilling all reporting requirements independently regardless of which test organization owns the personnel. The second issue is the independence of operational test data necessary to evaluate military suitability. A single joint contractor, development, and operational test team should be able to answer all management concerns. The contractor needs information to ensure contract delivery requirements are met, the development office needs information to adequately manage risk, and operational management needs information to report on system military suitability.

To fulfill these requirements while also reducing manpower and budgets, two test policy changes are needed. The first change allows use of data obtained during a contractor's development test program to fulfill military suitability evaluations in addition to operational certification readiness. The second proposed change will allow development and contractor test personnel to supplement operational test teams. Operational test interests are then addressed while also minimizing the test execution manpower needs. These two proposals would produce more streamlined and efficient test programs. The JDAM program reported benefits of conducting tests under operational conditions with production hardware:

> We required the contractor to demonstrate the key production processes in a production setting early in the competition. This enabled us to have confidence in what was being proposed. In the past we had demonstrations, but in a laboratory environment. Later we would find it was different in the production environment. We were hit with surprises that cost time and money.[14]

However, there are serious concerns which affect implementation of multiple-use data and combined teams.

One concern is using contractor test data to satisfy contractor, development, and operational test requirements. As mentioned earlier, contractor test data was unacceptably optimistic, especially for commercial acquisition programs.[15] Lessons emphasizing independent test to verify commercial claims resulted from optimistic data reports. Conversely, the F/A-18 E/F program met the reduced test requirements by using an integrated team approach.[16] Less testing was required because multiple test requirements were satisfied when all teams' test needs were considered in the planning process. Overall, test data efficiency was best when test data were gathered after the team was formed, and after the contract was awarded. Until commercial competitiveness

barriers are removed, concerns regarding data gathered solely by the contractor, to win a contract, will remain.

Further development and operational test streamlining is dependent upon well-integrated test teams. Combined teams rely on a pool of well-trained and experienced test personnel. However, both government and industry are questioning future access to specialized personnel.[17] The next chapter discusses test personnel issues and outline two initiatives to enhance teaming arrangements.

Notes

[1] Dr. Patricia. Sanders, Statement by the Director Test, Systems Engineering & Evaluation to the Subcommittee on Military Research & Development of the House National Security Committee, 27 February 1997, n.p.; on-line, Internet, 15 October 1997, available from http://www.acq.osd.mil/te/speeches/anders/testimony.htm.

[2] Jan Howell, "PACER CRAG — A COTS/NDI Lightning Bolt Program," briefing, Air Force Flight Test Center, Edwards AFB, CA., n.d.

[3] Ibid.

[4] William B. Scott, "Flight Test Jobs Shift Toward System Evaluation," *Aviation Week & Space Technology*, 12 June 1995, Vol. 142 Issue 24, CD-ROM. UMI-Proquest, 15 October 1997, 105.

[5] "The Integrated Test Team," *Navy Success Stories*, n.d., n.p.; on-line, Internet, 15 October 1997, available from http://www.acq-ref.navy.mil/jmcoms1.html.

[6] *Defense Acquisition Deskbook*, Version 2.2, Wright-Patterson Air Force Base, Ohio, ODUSD(AR and OUSD(A&T) API Deskbook Joint Program Office, 15 December 1997, n.p.; on-line, Internet, 12 January 1998, available from http://www.afam.wpafb.af.mil, AFMC - Lessons Learned fro Evaluating System Test Data, 7. Early Development of a Watch List for IOT&E.

[7] Ibid.

[8] Jan Howell, "PACER CRAG — A COTS/NDI Lightning Bolt Program," briefing, Air Force Flight Test Center, Edwards AFB, CA., n.d.

[9] *Title 10 — 10 Uniform Service Code; Armed Forces; Subtitle A — General Military Law Sec. 139 — Director of Operational Test and Evaluation; Sec.2399 — Operational Test and Evaluation of Defense Acquisition Programs*, n.d., n.p.; on-line, Internet, 22 October 1997, available from http://www.afam.af.mil

[10] DOD Directive 5000.2-R *Mandatory Procedures for Major Defense Acquisition Programs and Major Automated Information Systems Programs*, Change 2, 6 October 1997, n.p.; on-line, Internet, 4 February 1998, available from http://www.afam.wpafb.af.mil. Paragraph 3.4.3.

[11] AFI 99-101 *Development Test and Evaluation*, 1 November 1996, Paragraph 2.19 Operationally Relevant Testing.

Notes

[12] AFI 99-102 Operational *Test and Evaluation*, 22 July 1994, Paragraphs 1.3.1 and 1.3.2..

[13] *Defense Acquisition Deskbook,* Version 2.2, Wright-Patterson Air Force Base, Ohio, ODUSD(AR and OUSD(A&T) API Deskbook Joint Program Office, 15 December 1997, n.p.; on-line, Internet, 12 January 1998, available from http://www.afam.wpafb.af.mil, Lessons from the DOD IG-Test and Evaluation of Commercial Items. Specific references are: "Buying organizations must adequately test commercial products to ensure that the products live up to their specifications…" and "Buying organizations should consider requiring the performance of operational testing and evaluation as part of the source-selection process."

[14] "Air Force Joint Directed Attack Munition," n.d., n.p.; on-line, Internet, 24 September 1997, available from http://www.arnet.gov/References/Best_Pract_Docs/ AFJDAM.html.

[15] Ibid.

[16] Ibid.

[17] William B. Scott, "Flight Test Jobs Shift Toward System Evaluation," *Aviation Week & Space Technology,* 12 June 1995, Vol. 142 Issue 24, CD-ROM. UMI-Proquest, 15 October 1997, 105.

Chapter 4

Better, Faster, and Cheaper

> *History has proven, however, that new combat and commercial aircraft eventually will be necessary, if only to replace aging fleets that cannot be economically supported. At some point, both nations and companies will require well-trained, experienced flight test professionals to shepherd highly complex new aircraft through lengthy developments. But will they be available?*
>
> —William B. Scott
> *Aviation Week and Space Technology, 1995*[1]

Lessons from the previous two chapters discussed problems with test resources and test processes. The following chapter presents the third analysis area. The test personnel subcategories, test team training and team membership, accounted for the lowest percentage of lessons. Chapter 1 explained how data limitations prevented a full analysis of personnel effects on test programs. Even with limited information, data analysis attributed ten percent of lessons directly to personnel causes. Two personnel attributes affecting test programs were lack of test-trained and test-experienced manpower and inefficient use of existing manpower resources. Reducing the amount of manpower needed to plan and execute tests is one solution to the personnel problem. Automation of test planning and execution processes is a method that lessens manpower requirements. A second possible approach uses contractor personnel to execute government planning and execution processes.

Test Planning and Reporting Automation

Replacing Manpower with Technology

Limited test-experienced manpower affected both government and contractor teams. Data analysis showed several programs had optimistic test execution timelines that underestimated test and data reduction requirements. The PACER CRAG program, discussed in Chapter 2, is one example of optimistic test schedules.[2] Stringent timelines were optimized to produce a product in minimum time. In addition, test program extensions resulted from earlier planning problems, leaving the test programs unplanned and under-resourced. Although an efficient and well-trained test team would minimize schedule extensions, data analysis also showed test teams were understaffed and lacked test training. The lesson on Contractor Flight Test Engineering Capability states the effect of inadequately trained and experienced test personnel:

> The contractor's on-site flight test organization did not have an "in-house" engineering pool. Consequently, the "home office" engineer had to write the test plan, detailed procedure sheets, accomplish the test data reduction for engineering units, review data for instrumentation problems, write interim reports, and resolve technical problems. This impacted test planning, resulted in wasted flights, wasted money and clouded the true picture of the test program status. Because these "home officer" personnel had little test experience and had other responsibilities, the test plans were often incomplete or inaccurate which resulted in delays in writing the basic test plans. Frequently, data reduction took more than 60 days to complete.…These problems along with having to deal with communication delays…resulted in unnecessary testing, schedule slips, and wasted program funds. [3]

Reformed government DT&E test roles also added test program pressures. Government test execution agencies also experienced manpower reductions. The PACER CRAG program overextended DT&E test resources and the team struggled to support the test program.[4] As the government moved to smaller program offices, test

efficiency should have increased. Lessons showed overall program timelines were reduced. However, test timelines did not necessarily follow suit and in several cases the timelines lengthened. The FA-18 E/F program cited reduced timelines as part of their manpower reductions.[5] Conversely, the PACER CRAG test program actually increased from 40 planned test hours to 276 flight test hours at completion despite the moves to commercial practices.[6] Adaptive test planning, without additional manpower requirements, is a recommendation to address test program inconsistencies.

Adaptive Test Planning from the Beginning

Adaptive test planning automates test planning and execution processes. The approach minimizes manpower required to test plan, execute, and report while also providing a near-real-time induction of experience learned from other programs. The Defense Acquisition Deskbook is the first step in the automation process, but test tools automation must go further.

Test education and information transfer are complex because there are many information sources. One Air Force Materiel Command Internet site references an additional 278 related Internet sources for test information.[7] In addition, multiple CD-ROM test planning and execution software programs are available covering any aspect of the test process. A sample of software programs aiding the test process includes the Design Traceability Manager and Timeline Management Tool.[8] The Test Planning, Analysis, and Evaluation System is also available to aid test planning.[9]

In addition to test software programs, integrated test planning methods exist. Dr. Patricia Sanders, Director of Test, Systems Engineering & Evaluation for DOD, proposed

an integrated test planning method called the System Test and Evaluation Process (STEP):

> What is new with STEP is the interdependent way in which simulation and test tools are applied in support of the acquisition process.…The implications of STEP to the acquisition community are: Testers will be involved much earlier in the acquisition program.…Predictions of system performance can be made to assess the military worth of the system before any physical prototypes are built. Predictive simulations can be used to expand the scope of system testing and evaluate system performance in areas that are not readily testable. The models, simulations, and test data will have a much wider use throughout the system's life cycle.[10]

The STEP approach models the system first, simulates the expected performance, then physically tests the system, and finally iterates the test results back into the software model. Unfortunately, simulation capabilities are limited in their ability to handle software-intensive integrated systems. Several lessons reported simulation problems including random software test failures and difficulty representing complete systems using software models.[11]

Multiple test information sources, various test software tools, and simulation limitations increase test planning and execution complexity. Experienced and trained test personnel are absolutely necessary to efficiently plan and execute a test program in minimum time, but, future test organizations can anticipate fewer test personnel to accomplish test planning and execution.[12] Manpower resources must be used more efficiently. For the adaptive test approach to work, increased test automation is crucial. In addition, test automation should include the following:

1. Transfer knowledge from all DOD acquisition programs as quickly and efficiently as possible.
2. Accessible by both government and contractor personnel.
3. Use as the test planning standard by all agencies to maximize cross program utility.
4. Consolidate information sources to shorten research investments from program to program.

Reducing Manpower Requirements

Test Execution is Manpower Intensive

Adaptive test planning aids transfer of test experience and training across multiple programs and agencies. However, the transfer of information assumes the test team has the appropriate team membership to accomplish the program. From analysis, team membership accounted for two program lessons. The first lesson recommends a thorough analysis of known deficiencies of commercial equipment before contract award.[13] The second lesson recommends the government test execution team should take part in the request for proposal, sources selection, test planning working group, and test and evaluation master plan processes to ensure that test requirements are adequately stated.[14]

Both lessons learned mention SPO responsibility to address engineering and test requirements early in an acquisition cycle. In the first lesson, the SPO team either did not have sufficient manpower or experience levels needed to accomplish deficiency analysis. The second lesson emphasizes the need for additional test team membership from the beginning of a program. However, government manpower reductions will not be reversed and manpower levels may continue to decline. Congressional testimony by Dr. Patricia Sanders predicts continued personnel reductions:

> Test and Evaluation personnel will have been reduced approximately 35 percent by FY1998 from the peak level in the late 1980s. But we need to do better in this area.[15]

She further defines test execution manpower requirements, specifically related to open-air testing, as significant.

> … we find that open air testing accounts for only about 30 percent of the work we do - but it accounts for over 60 percent of our manpower.[16]

Government test team participation is the major means of increasing government test experience.[17] However, additional test personnel reductions will affect any future ability to adequately plan or execute test programs. A shared manpower approach is a concept that delegates limited test execution authority to contractor personnel. Air Force guidance outlines test responsibilities and requirements, however, contractor test execution methods must be validated for conformance to this guidance. Monitoring contractor knowledge and adherence to government test requirements, without added government manpower, are prerequisites to delegating test autonomy. The Federal Aviation Administration (FAA) already uses a similar delegated test approach.

Government Trusted Agent Concept

The FAA uses Designated Engineering Representatives (DERs) to minimize government manpower requirements while using contractors' processes to complete certification requirements.[18] The DER concept uses individuals appointed for their employer to approve, or recommend for FAA approval, company technical data gathered in the test process.[19] In some cases, the DER may personally evaluate and approve data. In other cases, a DER may ensure, through the company management system, the proper evaluation of technical data by other persons, then approve that data by certifying compliance with applicable regulations.[20]

DER certification requirements include demonstrated knowledge and experience working in specified engineering positions in addition to experience working directly with the FAA.[21] The certified personnel review test plans, approve test reports and provide flightworthiness certification data for commercial aviation aircraft.[22] Depending

upon the tasks required, the FAA also delegates approval authority for some test planning activities, allowing the contractor to complete test programs with minimal government involvement.[23] The FAA retains control of personnel certifications.[24] In effect, the contractor's engineering staff fulfills both government and industry's test requirements simultaneously. Best commercial practice, in this case, is actually a best government practice of using shared manpower resources.

The DER approach does have limitations, with the FAA reserving many functions for itself as the approval authority. Overall flight test plans, new methods or principles of testing, and new operational procedures approval are among the functions retained by the FAA.[25] The FAA still requires a staff of qualified and experienced test personnel. In addition, Air Force adoption of the DER approach may limit a program office's ability to complete operational test certification requirements. Laws and restrictions on independent verification of data, mentioned in Chapter 3, limit the proposal's usefulness. The DER concept will also require a staff to certify commercial engineers as government engineering representatives. The certifying staff may create a short-term manpower increase over existing levels. Overall, the DER concept's advantages outweigh the disadvantages. The concept does hold promise to advance commercial aviation's understanding of government test planning, test execution, and test reporting functions.

Enhanced teaming arrangements follow best commercial practice, however lessons discussed throughout the last three chapters show that even the best test teams require training and experience to plan and execute test programs. As the government downsizes, access to future trained and experienced test personnel is still an issue.[26] How long before too few experienced personnel are left to employ test lessons learned?

Notes

[1] William B. Scott, "Flight Test Jobs Shift Toward System Evaluation," *Aviation Week & Space Technology,* 12 June 1995, Vol. 142 Issue 24, CD-ROM. UMI-Proquest, 15 October 1997, 105.

[2] Jan Howell, "PACER CRAG — A COTS/NDI Lightning Bolt Program," briefing, Air Force Flight Test Center, Edwards AFB, CA., n.d.

[3] *Defense Acquisition Deskbook,* Version 2.2, Wright-Patterson Air Force Base, Ohio, ODUSD(AR and OUSD(A&T) API Deskbook Joint Program Office, 15 December 1997, n.p.; on-line, Internet, 12 January 1998, available from http://www.afam.wpafb.af.mil, AFMC - T&E Lessons Learned -RFP Prep/Source Selections Lesson Learned, 13. Contractor Flight Test Engineering Capability.

[4] Jan Howell, "PACER CRAG — A COTS/NDI Lightning Bolt Program," briefing, Air Force Flight Test Center, Edwards AFB, CA., n.d.

[5] "The Integrated Test Team," *Navy Success Stories,* n.d., n.p.; on-line, Internet, 15 October 1997, available from http://www.acq-ref.navy.mil/jmcoms1.html.

[6] Michael E. Bonner and Michael J. Scheall, *PACER CRAG C/KC-135 Avionics Modernization Evaluation,* AFFTC-TR-97-01, Air Force Flight Test Center, Edwards Air Force Base, California, May 1997.

[7] Jeff Braun and Tony Rivera, *Test and Evaluation (T&E) WWW Links,* 24 June 1997, n.p.; on-line, Internet, 22 October 1997, available from http://www.afmc.wpafb.af.mil/organizations/HQ-AFMC/DO/dop/dtet/d_tlinks.htm.

[8] Design Traceability Manager and Timeline Management Tool, Version 2.2 and 2.5 respectively, CD-ROM, Air Force Armstrong Laboratory Crew Systems Directorate, n.d.

[9] Test Planning, Analysis, and Evaluation System, Beta Version 4.1, CD-ROM, Air Force Armstrong Laboratory, Jun 1997.

[10] Dr. Patricia Sanders, Statement by the Director Test, Systems Engineering & Evaluation to the Subcommittee on Military Research & Development of the House National Security Committee, 27 February 1997, n.p.; on-line, Internet, 15 October 1997, available from http://www.acq.osd.mil/te/speeches/anders/testimony.htm.

[11] *Defense Acquisition Deskbook,* Version 2.2, Wright-Patterson Air Force Base, Ohio, ODUSD(AR and OUSD(A&T) API Deskbook Joint Program Office, 15 December 1997, n.p.; on-line, Internet, 4 February 1998,. available from http://www.afam.wpafb.af.

[12] William B. Scott, "Flight Test Jobs Shift Toward System Evaluation," *Aviation Week & Space Technology,* 12 June 1995, Vol. 142 Issue 24, CD-ROM. UMI-Proquest, 15 October 1997, 105.

[13] *Defense Acquisition Deskbook,* Version 2.2, Wright-Patterson Air Force Base, Ohio, ODUSD(AR and OUSD(A&T) API Deskbook Joint Program Office, 15 December 1997, n.p.; on-line, Internet, 12 January 1998, available from http://www.afam.wpafb.af.mil, Lessons from the DOD IG-Test and Evaluation of Commercial Items.

[14] *Defense Acquisition Deskbook,* Version 2.2, Wright-Patterson Air Force Base, Ohio, ODUSD(AR and OUSD(A&T) API Deskbook Joint Program Office, 15 December 1997, n.p.; on-line, Internet, 4 February 1998,. available from http://www.afam.wpafb.af.mil, AFMC - Lessons Learned for Evaluating System Test Data, 4. Random Software Test Failures.

Notes

[15] Dr. Patricia Sanders, Statement by the Director Test, Systems Engineering & Evaluation to the Subcommittee on Military Research & Development of the House National Security Committee, 27 February 1997, n.p.; on-line, Internet, 15 October 1997, available from http://www.acq.osd.mil/te/speeches/anders/testimony.htm.

[16] Ibid.

[17] William B. Scott, "Flight Test Jobs Shift Toward System Evaluation," *Aviation Week & Space Technology,* 12 June 1995, Vol. 142 Issue 24, CD-ROM. UMI-Proquest, 15 October 1997, 105.

[18] 8110.37B *Designated Engineering Representatives (DER) Guidance Handbook,* Federal Aviation Administration, 12 November 1996, n.p.; on-line, Internet, 15 October 1997, available from http://www.fedworld.gov/pub/faa-oai/faa-oai.htm.

[19] Ibid.

[20] Ibid., Paragraph 7a and b.

[21] Ibid.

[22] Ibid.

[23] Ibid.

[24] Ibid.

[25] Ibid., Appendix 1 - Typical Limitations on DER Functions.

[26]. William B. Scott, "Flight Test Jobs Shift Toward System Evaluation," *Aviation Week & Space Technology,* 12 June 1995, Vol. 142 Issue 24, CD-ROM. UMI-Proquest, 15 October 1997, 105.

Chapter 5

Conclusions

Our bottom line is…field a superior capability, affordably and in less time than our potential adversaries.

—Honorable Paul G. Kaminski
Under Secretary of Defense for Acquisition and Technology, 1995[1]

Six Flight Test Concepts

In the last three chapters, test lessons analysis identified several problems in recent test programs. From analysis conclusions, six recommendations followed which proposed test program initiatives for further consideration. Lessons used to arrive at these conclusions cross-cut many programs. In addition to an acquisition program lesson database, two Air Force and one Navy flight test program best practices added comparison and contrasting information. Finally, one Air Force program, PACER CRAG, provided detailed DT&E lessons adding analysis depth.

Listed in Chapter 1, analysis categories suggested three analysis questions. Analysis results showed the test capabilities category, including test planning and test execution, was the greatest source of test lessons. Recommendations to improve flight test planning and execution were:

1. Develop a single commercial industry standard governing military aviation testing.

2. Emphasize greater simulation configuration control and technological simulation advances, with software test and avionics integration simulation the highest priorities.

In the processes category, test redundancy and test streamlining lessons described the second greatest sources of problems. Recommendations to minimize test redundancy and improve test program streamlining were:

1. Refine the Integrated Test Team Concept to combine contractor, development, and operational test requirements in addition to innovative test program teaming.
2. Streamline development and operational test data requirements and allow contractor and development test personnel to supplement operational test teams.

Finally, test personnel was the third analysis area. Two alternative test approaches needed to maximize use of trained and experienced test personnel were:

1. Automate test planning and execution functions to reduce manpower investments.
2. Modify the Federal Aviation Administration Designated Engineering Representative concept to increase contractor test autonomy.

Areas for Additional Research

The above recommendations are general concepts which require additional research before implementation. Three topics arise as areas for detailed study to translate the concepts into execution plans. A DT&E simulation study would allow isolation of avionics simulation priorities and technologies needed to advance our current capabilities. Contractor and government team membership research is needed before additional tailoring of training and test team recommendations. Finally, the DER concept requires feasibility and cost studies before a true assessment of the concept viability is known. Continued research and analysis, in these areas especially, will provide the beginning foundation of future innovative test programs.

Conclusion

Acquisition reform forced changes to longstanding DT&E methods including SPO manning reductions, streamlined source selections, and military guidance cancellations. Lessons from recent programs reflected trends which, if uncorrected, will limit and possibly reverse future reform gains. Stressed test planning and execution capabilities, duplicated test requirements, and under-utilized test manpower illustrate inefficiencies left in the DT&E process. However, correcting these inefficiencies will require difficult choices.

The proposed recommendations are concepts which indicate an even larger test trend. Test capabilities, specifically test planning and test execution, accounted for 62 percent of total lessons examined. Both Air Force and Navy programs reported test processes and personnel successes, but not test capabilities achievements. Test processes and test personnel reform changes built upon previous successful initiatives. Test capabilities, however, were radically altered by reform measures. Tomorrow's successes will be fewer and require expensive test capabilities investments to achieve.

Successfully overcoming DT&E challenges will determine future program viability. Adaptive test planning and execution capabilities are necessary to execute future programs with less manpower investments. Test redundancy and limits on test streamlining are unaffordable luxuries. Future availability of trained test personnel is uncertain. However, these challenges may be met with investments in improved test capabilities, removal of test streamlining constraints, and maximizing test team arrangements. Our nation's strength lies in our ability to field superior systems quickly.

Future military capability is dependent upon our foresight to sharpen DT&E methods today.

Notes

[1] Paul G. Kaminski, "Reinventing DOD Test & Evaluation," 3 October 1995, n.p.; on-line, Internet, 17 September 1997, available from http://www.acq.osd.mil/te/speeches/kaminski/itea.doc.

Appendix A

Analysis Definitions and Criteria

Three primary categories were used to assess lessons learned: capabilities, processes, and personnel. Each category was further divided into two subcategories reflecting themes in the lessons learned analysis. A subjective evaluation of the most likely lessons cause determined the subcategory.

Capabilities

Resource availability to accomplish test planning and execution was the first category. Resources included documentation, software, and hardware technical needs.

Test Planning

Errors or successes occurring in test planning due to standards, specifications or contractual guidance.

Test Execution

Lessons learned which occurred during test execution attributable to simulations, simulator equipment, ground test or flight test conduct. Also included was post test analysis when test execution was dependent upon previous valid test results.

Processes

Test processes which facilitated or hampered accomplishing test objectives were considered in this category. Policies and guidance allowing accomplishment of contractor, development and operational teat requirements were also included in the analysis.

Test Redundancy

Integration of contractor, development, and operational test requirements lessons which added or reduced test needs. Also included in this subcategory were repeated test requirements due to data analysis procedures or test data recording techniques.

Limitations to Further Streamlining

Lessons learned attributed to laws, guidance, and policies that promoted or limited the test team's ability to streamline test accomplishment or minimize program test requirements.

Personnel

Test training and experience levels of the personnel planning and executing tests were considered in the third category. Team composition included contractor and government organizations and all phases of the acquisition cycle.

Training

Addressed test and evaluation experience and training levels of both government and contractor team members. Technical training provided to operational test groups was also included. The team's technical ability in planning, executing, analysis, and reporting

on test activities from definition of requirements until completion of acquisition program were inclusive analysis areas.

Membership

Lessons learned attributed to test team membership when making decisions affecting the test program. Decisions made by the team may occur at any phase of the program from requirement definition through completion of the program.

Appendix B

Lessons Learned Detailed Listing

Table 2. Lessons Learned General Categories

Lesson Learned	**P**	**E**	**R**	**L**	**T**	**M**	**Totals**
1. DOD Inspector General -Test and Evaluation of Commercial Items	9	5	2			1	17
2. AFMC - T&E Lessons Learned -RFP Prep/Source Selection	6	3		4	2	1	16
3. AFMC - Lessons Learned for Detailed Test Planning	2	4			2		8
4. AFMC - Lessons Learned for Performing Pre-Test Activities		5	4				9
5. AFMC - Lessons Learned for Evaluating System Data		3	4	1			8
6. Navy - Manufacturing Processes with Sole Source Contractor				1			1
7. Navy - Gun MK 45 Performance Based Specifications			1				1
Totals	**17**	**20**	**15**	**2**	**4**	**2**	**60**

Legend:
P Planning
E Execution
R Redundancy
L Limitations to Further Streamlining
T Training
M Membership

Table 3. DOD IG - Test and Evaluation of Commercial Items Lessons Learned Subcategory Classifications

Lesson Learned	**Subcategory**
Performance of OT&E as part of the Source Selection Process	Planning
Use conformance testing of commercial products in the source selection process	Planning
Develop a sensible test program using previous manufacturer and Government test results	Redundancy
Tailor testing to proven technology	Execution
Tailor testing to address program risk areas	Planning
Organize testing to maximize use of the tests	Redundancy
Plan the conduct of operational testing as early as possible	Planning
Ensure that performance pass or fail criteria are clearly specified in the contract	Planning
Analyze known deficiencies of commercial equipment before contract award	Membership
Share test plans and results with contractors	Execution
Test organizations maintain on-site representation during test execution	Execution
Ensure newly delivered equipment in initially single-sited for testing purposes to avoid spreading any undetected problems	Execution
Test software acquisitions in each fielded configuration	Planning
Adequately test commercial products because commercial specifications tend to be optimistic	Planning

Notes:

1. Subcategory Definitions are explained in Appendix A.
2. Lessons Learned in the above table, and subsequent tables, are presented in order as they appeared in the Defense Acquisition Deskbook database.

Table 4. AFMC - T&E Lessons Learned - Prep/Source Selections

Lesson Learned	Subcategory
1. Testing Suitability of GFE in Vibroacoustic Environments	Planning
2. Off-Site Testing	Membership
3. Contractual Provisions for Safety Chase	Planning
4. Test, Analyze, and Fix Activities	Planning
5. Software Regression Testing	Execution
6. Inadequate Installation Instructions	Execution
7. Test Data for Logistics Supportability Evaluations	Execution
8. Concurrency in Major System Developments and Modifications	Redundancy
9. Software Independent Verification & Validation	Planning
10. Development and Approval of Acceptance Test Procedures (ATP)	Planning
11. Environmental Stress Screening (ESS) Tests	Planning
12. Operational Unit Involvement in Development Flight Testing	Redundancy
13. Contractor Flight Test Engineering Capability	Training
14. Contractor Provided (Type 1) Training	Training
15. Flight Qualification Testing for Aircraft Modifications	Redundancy
16. Transporting Failed Parts after Testing	Redundancy

Table 5. AFMC-Lessons Learned for Detailed Test Planning

Lesson Learned	Subcategory
1. Test Unique Hazard Analysis	Training
2. Emergency Procedures to Support System Tests	Planning
3. Inherent Testing Risks	Training
4. Planning for Flight Test Crew Qualification and Checkout	Planning
5. Review of Test Aircraft Operating Limitations	Execution
6. Test Munitions Recovery and Demolition Records	Execution
7. Improper Fuel used in Explosive Atmosphere Testing	Execution
8. Configuration Control Deficiency at Simulation Facility	Execution

Table 6. AFMC-Lessons Learned for Performing Pre-Test Activities

Lesson Learned	Subcategory
1. Coordinating Test Configuration Changes	Redundancy
2. Aircraft Operating Limitations (AOL) for Test Flights	Execution
3. Incomplete Documentation of Test Flight Restrictions	Execution
4. Review of Aircraft Operating Limitations (AOL)	Execution
5. Comparing Model Characteristics to Test Aircraft Performance	Execution
6. Extensive Premature Coding of Software	Redundancy
7. Predicting Flight Failures from Ground Test Results	Execution
8. Offsite Flight Readiness Reviews	Redundancy
9. C-130 Autopilot Replacement Program (ARP) Flight Test	Redundancy

Table 7. AFMC - Lessons Learned for Evaluating System Test Data

Lesson Learned	Subcategory
1. Evaluating Reliability Test Results	Redundancy
2. Testing Real Gas Effects Against Predicted Aerodynamics	Execution
3. Analysis of Flight Test Data	Redundancy
4. Random Software Test Failures	Execution
5. Flight Test Data Analysis	Execution
6. Identification of OT&E Problems/Fix Cycle	Limitations to Streamlining
7. Early Development of a Watch List for IOT&E	Redundancy
8. Forwarding Service Reports	Redundancy

Table 8. Other Lessons Learned

Lesson Learned	Subcategory
1. Navy - Manufacturing Processes with Sole Source Contractor	Limitations to Streamlining
2. Navy - MK 45 Performance Based Specifications	Redundancy

Glossary

ACSC	Air Command and Staff College
AF	Air Force
AFFTC	Air Force Flight Test Center
AFI	Air Force Instruction
AFMC	Air Force Materiel Command
AFOTEC	Air Force Operational Test and Evaluation Center
AU	Air University
DAD	Defense Acquisition Deskbook
DER	Designated Engineering Representative
DOD	Department of Defense
DODD	Department of Defense Directive
DT&E	Development Test and Evaluation
FAR	Federal Acquisition Regulation
HQ	Headquarters
IG	Inspector General
IOT&E	Initial Operational Test and Evaluation
ITT	Integrated Test Team
JDAM	Joint Direct Attack Munition
JPATS	Joint Primary Aircraft Training System
MNS	Mission Need Statement
ORD	Operational Requirements Document
OT	Operational Test
OT&E	Operational Test and Evaluation
SPO	System Program Office
STEP	System Test and Evaluation Process
T&E	Test and Evaluation
USAF	United States Air Force

WWW	World Wide Web

Bibliography

Air Force Instruction (AFI) 99-101. *Development Test and Evaluation*. 1 November 1996.

AFI 99-102 Operational *Test and Evaluation*. 22 July 1994.

“Air Force Joint Direct Attack Munition.” N.d. [no date], n.p. [no pagination]. On-line. Internet, 24 September 1997. Available from http://www.arnet.gov/References/ Best_Pract_Docs/ AFJDAM.html.

Bonner, Michael E., and Scheall, Michael J. *PACER CRAG C/KC-135 Avionics Modernization Evaluation.* AFFTC-TR-97-01. Air Force Flight Test Center., Edwards Air Force Base, California. May 1997.

Braun, Jeff and Rivera, Tony. *Test and Evaluation (T&E) WWW Links.* 24 June 1997, n.p. On-line. Internet, 22 October 1997. Available from http://www.afmc.wpafb.af.mil/organizations/HQ-AFMC/DO/dop/dtet/d_tlinks.htm.

Defense Acquisition Deskbook. Version 2.2. Wright-Patterson Air Force Base, Ohio, ODUSD(AR and OUSD(A&T) API Deskbook Joint Program Office. 15 December 1997, n.p. On-line. Internet, 12 January 1998. Available from http://www.afam.wpafb.af.mil.

Design Traceability Manager and Timeline Management Tool. Version 2.2 and 2.5 respectively. CD-ROM. Air Force Armstrong Laboratory Crew Systems Directorate. N.d.

DOD Directive 5000.2-R *Mandatory Procedures for Major Defense Acquisition Programs and Major Automated Information Systems.Programs.* Change 2. 6 October 1997, n.p. On-line. Internet, 4 February 1998. Available from http://www.afam.wpafb.af.mil.

FAR Part 12. *Acquisition of Commercial Items.* 9 February 1998, 10.

8110.37B. *Designated Engineering Representatives (DER) Guidance Handbook.* Federal Aviation Administration, 12 November 1996. On -line. Internet, 15 October 1997. Available from http://www.fedworld.gov/pub/faa-oai/faa-oai.htm.

Howell, Jan. “PACER CRAG—A COTS/NDI Lightning Bolt Program.” Briefing. Air Force Flight Test Center, Edwards AFB, CA. N.d.

“The Integrated Test Team.” *Navy Success Stories.* N.d., n.p. On-line. Internet, 15 October 1997. Available from http://www.acq-ref.navy.mil/jmcoms1.html.

“Air Force Joint Directed Attack Munition.” N.d., n.p. On-line. Internet, 24 September 1997. Available from http://www.arnet.gov/References/Best_Pract_Docs/ AFJDAM.html.

“Joint Primary Aircraft Training System.” *Acquisition Reform Stories.* N.d., n.p. On-line. Internet, 15 October 1997. Available from http://www.safaq.hq.af.mil/acq-ref/stories/stories.htm.

Kaminski, Paul G. "Reinventing DOD Test & Evaluation." 3 October 1995, n.p. On-line. Internet, 17 September 1997. Available from http://www.acq.osd.mil/te/speeches/kaminski/itea.doc.

Lightning Bolt Update #14. SAF/AQ, Pentagon, VA. 6 August 1997, n.p. On-line. Internet, 21 January 1998. Available from http://www.safaq.hq.af.mil/acq_ref/bolts/update14.html.

OMB Circular A-119 *Federal Participation in the Development and Use of Voluntary Standards*. 20 October 1993, n.p. On-line. Internet, 4 February 1998. Available from http://www.afam.wpafb.af.mil.

Perry, William. "Mandate for Change." Briefing presented to the House Armed Services Committee and the Governmental Affairs Committee. February 1994, 18. On-line. Internet, 9 March 1998. Available from http://www.acq.osd.mil/ar/#library/mand24.pdf.

Sanders, Patricia. "Statement by the Director Test, Systems Engineering & Evaluation to the Subcommittee on Military Research & Development of the House National Security Committee." 27 February 1997, n.p. On-line. Internet, 15 October 1997. Available from http://www.acq.osd.mil/te/speeches/sanders/testimony.htm.

Scott, William B. "Flight Test Jobs Shift Toward System Evaluation." *Aviation Week & Space Technology,* 12 June 1995, Vol. 142 Issue 24, 105. CD-ROM. UMI-Proquest, 15 October 1997.

Scott, William B. "New Global Pressures Reshape Flight Testing." *Aviation Week & Space Technology,* 12 June 1995, Vol. 142 Issue 24, 62. CD-ROM. UMI-Proquest, 15 October 1997.

Soler, Oscar. JDAM, Air Force. Best Practices Documents. AFJDAM, n.d. On-line. Internet, 24 September 1997. Available from http://www.arnet.gov/References/Best_Pract_Docs/AFJDAM.html.

Strategic Technologies, Inc. *Independent Assessment of Operational Flight Programs Modification/Update Process.* Albuquerque, NM: Strategic Technologies, Inc., n.d.

Test Planning, Analysis, and Evaluation System. Beta Version 4.1. CD-ROM. Air Force Armstrong Laboratory, Jun 1997.

Title 10—10 Uniform Service Code; Armed Forces; Subtitle A — General Military Law Sec. 139 — Director of Operational Test and Evaluation; Sec.2399 — Operational Test and Evaluation of Defense Acquisition Programs. N.d., n.p. On-line. Internet, 22 October 1997. Available from http://www.afam.af.mil.

Air Command and Staff College
Maxwell AFB, Al 36112